A wavelet-based approach for large wind power ramp characterisation

Cristóbal Gallego[1], Alexandre Costa[2], Álvaro Cuerva[1], Lars Landberg[3], Beatrice Greaves[4] and Jonathan Collins[4]

[1] ETSI.Aeronáuticos, Universidad Politécnica de Madrid, 28040 Madrid, Spain
[2] Wind Energy Unit, CIEMAT, 28026 Madrid, Spain
[3] Short Term Forecasting Department, GL Garrad Hassan, Copenhagen, Denmark
[4] Short Term Forecasting Department, GL Garrad Hassan, Bristol, UK

KEYWORDS

ramp; wind power; wavelets; short-term forecast, electricity market

Correspondence

C. Gallego, DVA-ETSI.Aeronáuticos, Universidad Politécnica de Madrid, Pza. Cardenal Cisneros 3, 28040 Madrid, Spain.
E-mail: cristobaljose.gallego@upm.es

NOMENCLATURE

ΔP	Variation of power output
P_R	Rated power
Δt	Duration
t_0	Timing
I_t	Indicator function
S_t	Criterion function
S_0	Threshold value
ΔP_0	Power variation threshold
f_r	Relative frequency
$y(t)$	Continuous signal
λ	Time scale
τ	Time in the time-frequency domain
$\psi(t)$	Mother wavelet
$\psi^{\tau,\lambda}(t)$	Shifted and scaled wavelet (continuous)
$\{y_t\}$	Discrete time series
$W^{\tau,\lambda}$	Wavelet transform coefficients
$\psi_t^{\tau,\lambda}$	Shifted and scaled wavelet (discrete)
$W_H^{\tau,\lambda}$	Wavelet transform based on the Haar family wavelet
$\{R_t\}$	Ramp function
$\{r_t\}$	Relative ramp function
$\{r_t^u\}$	Relative ramp function during ramp-up events
$\{r_t^d\}$	Relative ramp function during ramp-down events
$\{r_t^0\}$	Relative ramp function during non-ramp events
$\{g_t^a\}$	Gradient function of order a
$\{f_t^\lambda\}$	Filtered function
$\{r_*^u\}$	$\{r_t^u\}$ sorted in decreasing order
$\{r_*^d\}$	$\{r_t^d\}$ sorted in decreasing order
$\{r_*^0\}$	$\{r_t^0\}$ sorted in decreasing order
$\{r_{t:h}^u\}$	Values of $\{r_t^u\}$ occurred at time h

$\{r^d_{t:h}\}$	Values of $\{r^d_t\}$ occurred at time h
$\{r^0_{t:h}\}$	Values of $\{r^0_t\}$ occurred at time h
$\{r^u_{*:h}\}$	Values of $\{r^u_{t:h}\}$ sorted in decreasing order
$\{r^d_{*:h}\}$	Values of $\{r^d_{t:h}\}$ sorted in decreasing order
$\{r^0_{*:h}\}$	Values of $\{r^0_{t:h}\}$ sorted in decreasing order
$C_{ij}(l)$	Cross-correlation function of ramp functions i and j for lag l
φ_i	Autoregressive coefficients of autoregressive model
k	Prediction horizon
$\hat{y}_{t+k\|t}$	Forecast of y_{t+k} provided at time t
MSE	Mean squared error
MSE^u	Mean squared error during ramp-up events
MSE^d	Mean squared error during ramp-down events
MSE^0	Mean squared error during non-ramp events

1. INTRODUCTION

Wind energy is one of the most important renewable resources. The level of installed wind capacity has increased significantly over the past decade, and many countries have already reached high levels of wind power penetration (between 5% and 30%[1]). However, both the variability and the uncertainty of the wind represent certain drawbacks compared with traditional energy sources, leading, e.g., to higher ancillary service requirements to maintain the electrical system balance. In fact, the integration of high wind power amounts has been conditioned to specific actions to reduce the impact of the mentioned drawbacks, such as the implementation of accurate wind power forecasting models, especially within the short-term context.[2–5]

The increasing size of wind farms has emphasised the effect of ramp events, i.e. large variations of the wind power output of a certain wind farm or a portfolio within a short period (up to a few hours). From a technical point of view, ramp events may be perceived by the Transmission System Operators (TSO) as critical perturbations in the generation side, especially in small balancing areas.[3] From an economic point of view, energy traders of liberalised electricity markets incur penalties because of deviations in scheduled energy. In both cases, forecasting wind power ramp events could represent a means to mitigate these problems, contributing to large-scale wind power integration.[6–8] The booming importance of wind power forecasting during extreme events is observed in the increasing number of studies and projects on this topic, with the European SafeWind Project being one of the most representative of this line of research.[9] However, it is important to note that most of the publications still belong to proceedings, conferences, and enterprise technical reports because of the novelty of this topic (see Thiesen,[10] Collier *et al.*,[11] Zack *et al.*[12] and Girard *et al.*,[13] among others). As a consequence of the recent activity, some considerations related to ramp forecasting are beginning to be highlighted. For example, a deeper classification of the causes of ramp events is still needed.[14,15] Ramp events can be motivated by a broad number of meteorological processes that occur at different time/spatial scales from horizontal processes (e.g. frontal systems in the synoptic scale, thunderstorms and hurricanes in the mesoscale) to vertical processes (e.g. dry convection).[15,16] Ramp events may also be conditioned by microscale effects such as wake effects and yaw misalignment, as well as other features such as the non-linearity of the wind turbine power curve, the cut-out wind speed and grid faults (power dips). Hence, a variety of ramp event causes must be considered to address wind power ramp forecasting. In the case of an oriented ramp event forecasting tool fed with the outputs of a General Circulation Model, the performance obtained would be limited by its resolution because mainly large-scale meteorological processes (synoptic processes) would be expected to be modelled, ignoring ramp events caused by mesoscale and microscale effects. Instead, the use of mesoscale models permits the modelling of meteorological processes with shorter life cycles, which usually leads to reductions in the amplitude of the forecasting error. However, the so-called phase error (which is observed when a certain process occurs earlier or later than forecasted) does not necessarily diminish with an increment in resolution because it may come from initial and boundary conditions provided by General Circulation Models.[17,18] Hence, approaching the problem of wind power ramp forecasting requires a well-defined strategy and a clear idea of the nature of the ramps intended to be modelled. Another recent question related to ramp forecasting is the criterion used to tune forecasting models. Traditional wind power prediction tools are currently based on the minimisation of a global error criterion (e.g. the root mean squared error), which severely penalises the mentioned phase errors during ramp events. As a result, the global minimisation approach tends to predict smoother ramps than the observed ones so as to cover a broader time window. Therefore, specific performance criteria focusing on ramp event forecasting have been demanded.[7,14,15] Finally, most of the related analyses agree on the lack of a robust ramp definition.[14,15,19,20] Intuitively, a ramp event is a large variation in wind power output within a short period. Ramps have usually been defined in a binary way by fixing two threshold values related to the absolute power variation and the duration of the event observed in the wind power time series. These thresholds are usually defined ad hoc, which has led

to several different ramp definitions.[7,14,15,19–21] These choices are understandable because the notion of ramp event may vary depending on end-user requirements, but a well-defined methodology to relate the mentioned thresholds to end-user needs is usually missing. In addition, several studies point out that a binary definition may artificially lead to the reduced performance of a ramp forecasting model.[7,19]

This work addresses the problem of characterising ramp events in wind power time series on the basis that the binary definition usually adopted is likely to be affected by several shortcomings. An innovative methodology for ramp event characterisation based on the wavelet transform (WT) is introduced. Only one parameter must be assessed from expertise as an input (related with the maximum duration of ramps). As a result, the ramp function is obtained, which provides a continuous index related to the ramp intensity at each time step of a given wind power time series.

This article is organised as follows: the binary definition of a ramp event is introduced and discussed in Section 2. In Section 3, the WT is introduced and the ramp function is defined. Section 4 collects several examples where the ramp function can be employed to characterise some features of the ramp performance of a wind farm. Finally, Section 5 gathers and discusses the main results obtained in this study.

2. RAMP EVENT GENERALITIES

This section addresses the notion of a ramp event. First, the main features of a ramp event are described. Next, the use of the indicator function as a basis for the binary ramp classification is introduced, as well as several specific cases presented in recent works. Finally, the implications of the use of the binary definition are outlined.

The term *ramp event* makes reference to a large variation in wind power output that is observed on a wind farm (or in a portfolio) within a short period (up to a few hours). To characterise a ramp event, the following set of parameters is usually employed:[7,11,22]

- Magnitude (ΔP): the variation of the power output during the event. ΔP is commonly defined as a percentage of the rated power, P_R.
- Duration (Δt): the period in which a large variation takes place.
- Timing (t_0): a time instant that specifies the time of the ramp event. t_0 can be defined as the starting time of the large variation, although the central time of the event can also be employed.

Another parameter of interest is the ramp rate, which is derived from the previous variables ($\Delta P/\Delta t$) and provides an idea of the ramp intensity. In addition, the ramp direction is usually also considered, i.e. whether the ramp event represents an increment or a decrement in power.

To assess these parameters during a specific extreme event, the ramp needs to be previously identified. In some cases, a ramp event can be intuitively recognised from the wind power time series under analysis. Figure 1 (top) clearly illustrates two consecutive ramp events observed during 25 and 27 January in the wind farm of Ariz, located in a region in the northwest of Spain that has a complex terrain. The wind farm of Ariz consists of 32 wind turbines with a rated power of 19.2 MW. The time resolution of the wind power time series is 1 h. In this case, variations in the $\Delta P_{25\text{th}} = 91.1\%$ and $\Delta P_{27\text{th}} = 83.3\%$ of the rated power were observed within a duration of $\Delta t = 4$ h. On the other hand, Figure 1 (bottom) shows a period where high variations in power were experienced (between 11 and 13 December) such that the non-monotonic evolution of the time series makes it difficult to assess the ramp magnitude without having to make an arbitrary decision related to the ramp duration. This example illustrates that identifying ramps *with the naked eye* may be a non-trivial task.

To deal with the subjective concept of ramp, all of the studies reviewed relied on the definition of specific thresholds to decide whether a certain event could be considered a ramp in a binary manner. The binary classification is based on the indicator function, I_t, defined as follows:

$$I_t = \begin{cases} 1, & \text{if } S_t \geq S_0 \\ 0, & \text{if } S_t < S_0 \end{cases} \tag{1}$$

where S_t is a predefined criterion function, $S : \mathbb{R}^n \longrightarrow \mathbb{R}$, evaluated at time t and S_0 represents a threshold. A specific event occurs during a period that causes the criterion function to go above threshold (i.e. $I_t = 1$). The indicator function was used in Hedley and Keffer[23] to decide whether fluid motion could be considered turbulent. However, it can be used for other purposes simply by considering the appropriate criterion function S_t.

The indicator function allows the identification of ramp events in a wind power time series by considering a certain criterion function along with a threshold. In most of the studies reviewed, S_t is defined at time t on the basis of the largest variation of the power observed during a period Δt that includes t. In other words, P_t, the power output at time t, is considered to belong to a ramp event if t belongs to any of the periods Δt in which a variation of power ΔP goes beyond a certain value ΔP_0. The sign of the variation determines whether the ramp is positive ($I_t = 1$) or negative ($I_t = -1$).

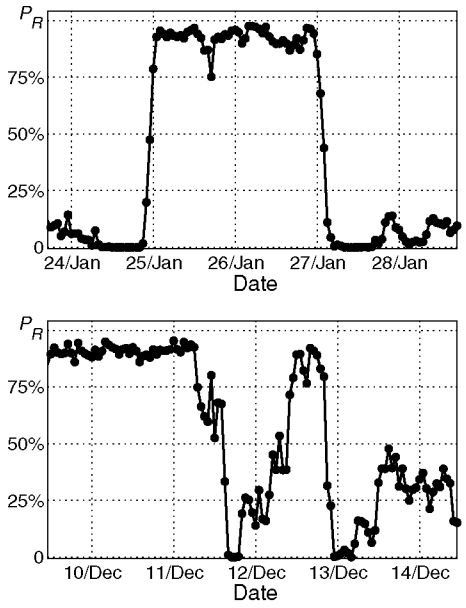

Figure 1. Examples of ramp events. A period with two clear large ramp events (top) and a period where ramp events are barely identifiable (bottom).

Mathematically,

$$I_t = \begin{cases} 1, & \text{if} \quad (P_{t_1+\Delta t} - P_{t_1}) \geq \Delta P_0 \quad \text{with } t \in [t_1, t_1 + \Delta t] \\ -1, & \text{if} \quad (P_{t_1+\Delta t} - P_{t_1}) \leq -\Delta P_0 \quad \text{with } t \in [t_1, t_1 + \Delta t] \\ 0, & \text{otherwise} \end{cases} \tag{2}$$

The indicator function of a certain wind power time series can be readily obtained by considering a set of two parameters $(\Delta t, \Delta P_0)$. The threshold ΔP_0 is usually given as a percentage of the rated power of a wind farm, $\%P_R$, although other criteria can be applied. For example, Gallego *et al.*[24] defined ΔP_0 as a percentage of σ_g, the standard deviation of the first-differenced wind power time series.* Bossavy *et al.*[21] proposed a binary classification based on variations experienced by a filtered signal $\{p_t^f\}$ instead of the original wind power time series. The filtered signal is obtained from a low-pass filter where a parameter related to the number of averaged measures, n, must be set. This parameter permits to tune the sensitivity to short period variations in the original time series (i.e. the characteristic time lengths of the ramps that are considered of interest). Setting the value of n replaces the election of the parameter Δt employed in previous definitions. As a result, $(n, \Delta P_0)$ provides a binary classification ramp/non-ramp for each time step of the original wind power time series.

Table I summarises the thresholds reviewed.

2.1. Implications of the use of the ramp binary definition

As mentioned previously, the indicator function is generally employed to identify ramp events within wind power time series. Its applicability is due to its simplicity given that only two parameters need to be assessed. It is reasonable to think that these two parameters should be set by the end-user. A similar idea was pointed out in Pinson[25] when the Safewind project partners were asked for their perception of an extreme event. Consequently, different ramp definitions may arise for each specific case depending on the size of the wind farm/portfolio and the cost function considered (e.g. costs of ancillary services and penalties of a certain electricity market). In addition, the parameters mentioned are commonly related to thresholds in power variations and periods, which ease the labour of defining ramps to end-users who may not be familiar with complex analysis tools. It is important to remark that the way in which ramps are identified may have appreciable impact on the activity of the forecaster/modellers because it highlights those periods where meteorological processes and

*The first-difference of a certain time series $\{P_t\}$ is defined as $\Delta P_t = P_t - P_{t-1}$.

Table I. Definitions of ramp events reviewed in recent studies.

	ΔP_0	Δt	Kind of ramp	Study case size
Cutler *et al.*[14]	75% P_R	3 h	Hourly resolution (slow ramps)	65 MW
Cutler *et al.*[14]	65% P_R	1 h	10 min resolution (fast ramps)	65 MW
Truewind[15]	20% P_R	1 h	Ramp-up	Not specified (portfolio)
Truewind[15]	15% P_R	1 h	Ramp-down	Not specified (portfolio)
Potter *et al.*[7]	10% P_R	1 h	No distinction	$\sim$ GW
Greaves *et al.*[19], Collier *et al.*[11]	50% P_R	4 h	No distinction	3–240 MW
Bradford *et al.*[20]	20% P_R	1 h	No distinction	Not specified (wind farm and portfolio)
Bossavy *et al.*[21]	50% P_R	n/a	No distinction	Not specified (wind farm)
Gallego *et al.*[24]	σ_g	1 h	No distinction	33 MW

P_R stands for the rated power of a wind farm, and σ_g is the standard deviation of the first-difference of the wind power time series.

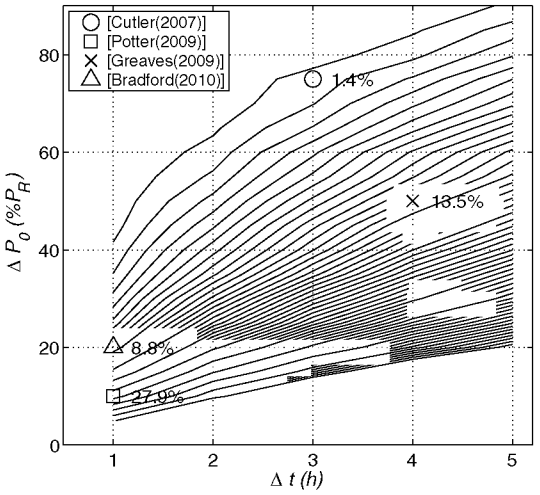

Figure 2. Relative frequency of ramp events, $f_r(\%)$, as a function of different thresholds $(\Delta t, \Delta P_0)$ for the power output time series of Ariz during a year. ΔP_0 stands for the wind power variation as a percentage of the rated power $\%P_R$. Δt stands for the maximum duration in which the power variation takes place. Contour lines interpolate constant relative frequencies in steps of 1%. The $f_r(\%)$ obtained with some of the reviewed definitions are also shown.

operational states should be carefully regarded to analyse the ramp causes. However, the use of the indicator function as a basis for ramp identification entails several shortcomings.

First, the number of ramp events observed in a wind power time series may become sensitive with respect to fixed threshold values. To illustrate this fact, Figure 2 shows the relative frequency of ramp events in Ariz during 1 year period, as a function of the thresholds considered. The relative frequency f_r is the percentage of times in which ramps was observed:

$$f_r(\%) = 100 \cdot \frac{\sum_{t=1}^{N} |I_t|}{N} \tag{3}$$

where N is the number of elements of the wind power time series. The specific case of four reviewed definitions has been pointed out. Contour lines interpolate constant relative frequencies (in steps of 1%) to provide insights about the sensitivity of the relative frequency with respect to both parameters $(\Delta t, \Delta P_0)$. In addition, Table II illustrates the relative frequency obtained with different ramp definitions for the case of Ariz when the magnitude threshold ΔP_0 is modified by $\pm 5\% P_R$. For example, considering the threshold $\Delta P_0 = 45\% P_R$ instead of $50\% P_R$ (defined in Greaves *et al.*[19]) increases the percentage of times where ramps were observed by 19.3%. This issue could have been expected since the distribution of power gradients in wind power time series tends to be a smooth function of both the amplitude and the duration considered, as discussed in Greaves *et al.*[19] Along these lines, Cutler *et al.*[14] proposed a sensitivity analysis for the threshold assessment.

Second, the use of a binary definition implies that ramp events are identified with no distinctions on their characteristics (e.g. ramp intensity and duration). This fact places conditions on the approach to wind power ramp forecasting for both the end-users and the forecaster/modellers. The use of the binary definition constrains end-users to translate their cost function into a binary perspective so as to identify what level of ramp intensity constitutes a problematic ramp event. Therefore,

Table II. Relative frequency of ramp events for different ramp definitions (column 1) and considering a modified amplitude threshold $\Delta P_0 \pm 5\% P_R$ (columns 2 and 3).

	$f_r(\Delta t, \Delta P_0)$	$f_r^+(\Delta t, \Delta P_0 + 5\% P_R)$	$f_r^-(\Delta t, \Delta P_0 - 5\% P_R)$	$100 \cdot \frac{f_r - f_r^+}{f_r}$ (%)	$100 \cdot \frac{f_r - f_r^-}{f_r}$ (%)
Cutler *et al.*[14]	1.4%	0.4%	1.9%	-71.4%	35.7%
Potter *et al.*[7]	27.9%	15.7%	49.5%	-43.7%	77.4%
Greaves *et al.*[19]	13.5%	10.9%	16.5%	-19.3%	22.2%
Bradford *et al.*[20]	8.8%	5.3%	15.7%	-39.7%	78.4%

Columns 4 and 5 show the relative variation between column 1 and columns 2 and 3, respectively.

defining a ramp as a variation of $\Delta P_0 = 50\% P_R$ within a period of $\Delta t = 4$ h implies that variations in $\Delta P = 90\% P_R$ have the same level of impact than the previous one. Furthermore, this definition implies that variations of $\Delta P = 49\% P_R$ within the same period have no impact at all. With respect to the forecaster/modellers, the ramp binary classification supports the notion that ramps are similar to one another even if ramps with different characteristics are usually observed. Few studies have proposed some degree of distinction by defining different thresholds depending on the direction of the ramp (ramp-up or ramp-down[15]) or the time series resolution (slow ramps and fast ramps[14]). In any case, the binary definition does not permit forecaster/modellers to exploit potential relationships between different ramp levels and explanatory variables that span a continuous range, such as outputs from Numerical Weather Prediction models or data from SCADA or off-site locations.

In summary, a suitable approach to wind power ramp forecasting should be able to deal with the fuzzy limits of the ramp notion to preserve more information than the binary classification. In this regard, it is probably more interesting to provide end-users and forecaster/modellers with tools oriented to better characterise the ramp performance of a wind power time series instead of conditioning the ramp definition to track optimal thresholds of a binary classification. This change would eventually permit more flexible and reliable actions from both end-users and forecasters.

3. THE RAMP FUNCTION

In this section, the ramp function is introduced and proposed as a means of characterising the ramp performance of a wind power time series. The ramp function is intended to overcome the drawbacks of the binary definitions mentioned in Section 2.1. The proposed function is defined using the WT, which is briefly described, and provides a continuous index of the intensity of the ramp at each time step.

3.1. The wavelet transform

Wavelet analysis is a relatively recent mathematical technique for signal analysis.[26] Wavelet analysis has become popular in the last few decades, especially because developments in computer science permit many different disciplines to take advantage on the properties of wavelets. One of the first applied studies on wavelets was related to seismic surveys.[27] Since then, wavelet analysis has been successfully applied with different purposes in a variety of fields[28] such as image processing,[29] edge detection,[30–32] time series forecasting[33] (electrical load forecasting,[34,35] electricity price forecasting[36]), chemical analysis[37] and tool condition monitoring.[38] The time-frequency decomposition of wavelet analysis has also been useful in the atmospheric sciences.[39] Specifically, several studies (see Szilagyi *et al.*,[40] Krusche and De Oliveira[41] and Watanabe[42] and the references therein) employed wavelet analysis to relate abrupt changes experienced by certain variables (temperature, wind speed or humidity) to coherent structures (large-scale turbulence processes in the atmospheric surface layer).

The WT is proposed to address certain aspects of signal processing that the Fourier transform is unable to deal with. Specifically, the frequency content of a certain signal $y(t)$ provided by the Fourier transform does not allow the analysis of events that take place during specific time windows, which are of interest for certain applications. In contrast, the WT permits to characterise the time-frequency content of a signal, providing information on both the location and the scale of different events. This feature is obtained because the transformation is based on functions (wavelets) that are localised in time and frequency. These wavelets are merely shifted and scaled versions of the so-called mother wavelet, $\psi(t)$, as given in the following equation:

$$\psi^{\tau,\lambda}(t) = \frac{1}{\sqrt{\lambda}} \psi\left(\frac{t-\tau}{\lambda}\right) \tag{4}$$

where τ relates to the shift and λ relates to the dilation.

The WT of a discrete time series $\{y_t\}$ consists of a set of coefficients obtained through the following equation:

$$W^{\tau,\lambda} = \sum_{-\infty}^{\infty} \psi_t^{\tau,\lambda} \cdot y_t \tag{5}$$

where $\tau \in \mathbb{Z}$, $\lambda \in \mathbb{Z}^+$ and $\psi_t^{\tau,\lambda}$ relates to the wavelet function employed.

Several wavelet functions with different mathematical properties exist (e.g. Daubechies or Meyer wavelets). The appropriate wavelet function usually depends on the purpose of the study. One of the most famous and simplest wavelets is the Haar wavelet. When the Haar wavelet is employed to perform a WT, information about the gradient experienced by the signal at different time scales is obtained.[43] For example, considering that $W_H(\tau, \lambda)$ is the WT obtained with the Haar wavelet, $W_H(\tau_0, \lambda_0)$ is related to the gradient experienced by the signal in $t = \tau_0$ within a time window of $\Delta t = \lambda_0$. Figure 3 illustrates $\psi_t^{\tau,\lambda}$ for the case of the Haar wavelet family for several scales λ. Each λ is related to a time scale $\Delta t = \lambda \cdot n$ hours, n being the number of records per hour in the time series (e.g. $n = 1$ for an hourly time series). Note that the Haar wavelet is defined in such a way that the sign of the coefficient obtained is the opposite of the sign of the gradient experienced by the time series.

A detailed description of the WT is beyond the scope of this work. For further theoretical details, several studies can be reviewed (see Mallat,[44] Daubechies,[26] Percival and Walden,[43] Burrus *et al.*,[45] Aboudafel and Schlicker[46] and Boggess and Narcowich,[47] among others).

3.2. Definition of the ramp function

Ramp events represent local events in a wind power time series and are characterised by sharp variations in power. Despite the fact that the magnitude and duration may vary from one ramp to another, the essential concept of a ramp is that a certain large gradient is maintained during consecutive time steps of the time series. Hence, the ramp pattern can be said to show self-similarity, which means that the whole event (wind power ramp) is similar to a smaller part of it (i.e. the shape is preserved under different scales). The use of different scales to characterise abrupt changes within signals has been widely employed; Canny[30] is one of the most well-known studies, where an optimal filter oriented to edge detection was proposed and applied to image processing. Mallat and Hwang[32] proved that the WT permits the detection of singularities in a signal from local maximum coefficients across different scales. We therefore propose the use of the WT for detecting and characterising ramp events within wind power time series. The use of the Haar function is proposed since we are interested in revealing local events where a high gradient is observed within a range of time scale, although further research could address the suitability of different wavelets.

Let $W_H^{\tau,\lambda}$ be the WT based on the Haar wavelet of a certain wind power output time series $\{y_t\}$. For the sake of simplicity, the conventional Haar wavelet was replaced by its additive inverse (by changing the sign of $\psi_t^{\tau,\lambda}$) to obtain positive coefficients for positive gradients of the time series.

Considering a period where a ramp event is observed, during this period the WT $W_H^{\tau,\lambda}$ would provide increasing coefficients for a broad range of time scales λ, since each $W_H^{\tau,\lambda}$ is related to the gradient experienced at time $t = \tau$ evaluated in a time scale $\Delta t = \lambda$. The increasing value of $W_H^{\tau,\lambda}$ with respect to λ is because the scale contributes as $\lambda^{-1/2}$ in equation (4). Hence, the mentioned self-similarity used to characterise the notion of a ramp event (i.e. the experienced gradient is high for both short and large scales, where 'large' means close to the characteristic time of the ramp event) can be easily identified by intense vertical lines in Figure 4 (right). Next, considering a period where $\{y_t\}$ experiences high-frequency fluctuations (i.e. positive and negative consecutive variations), similar coefficients are obtained for short time scales although the effect mentioned previously is not observed for large time scales. This effect is due to smoothing from evaluating the gradient

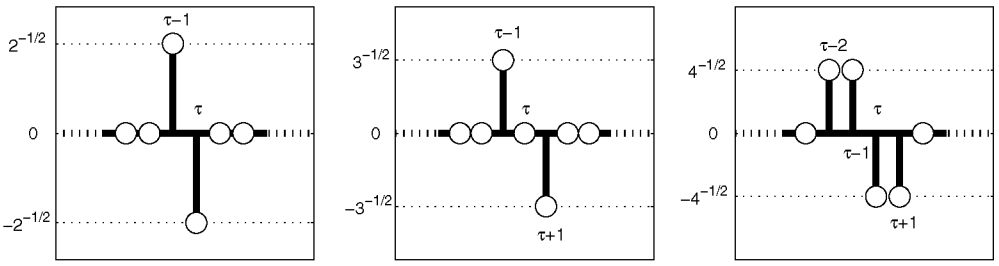

Figure 3. $\psi_t^{\tau,\lambda}$ for the case of the Haar wavelet family, scales $\lambda = 2$ (left), $\lambda = 3$ (centre) and $\lambda = 4$ (right).

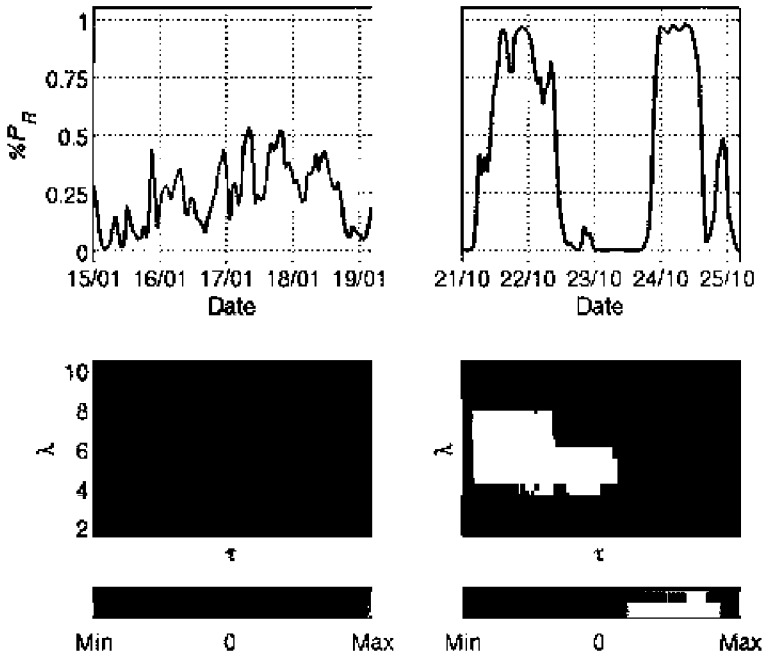

Figure 4. Coefficients of the wavelet transform based on the Haar function, $W_H(\tau, \lambda)$, during two different situations: a fluctuation period (left) and a ramp event period (right).

within a larger period, which highlights the fact that the net variation of the power during the fluctuations is small. This effect can be observed in Figure 4 (left). Finally, the coefficients $W_H^{\tau,\lambda}$ related to those periods where the time series do not show high gradients (large ramps or fluctuation periods) are expected to be close to zero for every scale λ considered.

The labour of characterising a ramp can take advantage of previous considerations in a simple way by defining the ramp function, $\{R_t\}$, as the addition of the wavelet coefficients $W_H^{\tau,\lambda}$ at time $t = \tau$ for the interval of scales λ given by $[\lambda_1, \lambda_N]$:

$$R_t(\lambda_1, \lambda_N) = \sum_{\lambda=\lambda_1}^{\lambda_N} W_H^{t,\lambda} \tag{6}$$

By performing this computation, $\{R_t\}$ becomes related to the sharpness of the ramp events because it gathers at each time step the contribution of the gradient evaluated under different time scales.

A crucial question is which time scales λ should be considered. The time scale λ_1 is the smallest scale to be considered. In the case of the WT of a discrete time series, the minimum possible value is $\lambda_1 = 2$ because at least two values of $\{y_t\}$ are required to evaluate a gradient. However, λ_N, the highest scale to be considered, must be established on the basis of expert knowledge. As mentioned previously, $W_H(\tau_0, \lambda_0)$ provides information about the gradient experienced by the time series in $t = \tau_0$ within a time window of $\Delta t = \lambda_0$. Hence, λ_N is the maximum time window where a local gradient is to be evaluated. The labour of characterising ramp events within a wind power time series requires λ_N to be related to the maximum duration expected for a ramp event (in terms of time steps). Choosing a low value for λ_N may result in a lower performance of the ramp function when differentiating between large ramp events and fluctuation periods. Conversely, values for λ_N related to very large time scales may cause a mixing of the gradients of different consecutive ramp events. Hence, on the basis of the analysis carried out with the available wind power time series during this work, we suggest that appropriate values for λ_N may be in the range of 5 to 10 (for hourly time series). The impact of this parameter on the ramp characterisation is addressed in depth in Section 3.3.

However, it should be noted that the proposed ramp function is not expected to be highly sensitive to small variations in λ_N, mainly because R_t is built from the addition of the WT coefficients across scales from λ_1 to λ_N (see equation (6)), which reduces the potential impact of considering an additional scale λ_{N+1}. For example, the relative frequency of ramp events* for the case of the wind farm of Ariz ranged from 10.0% (case of $\lambda_N = 5$) to 11.4% (case of $\lambda_N = 8$), which

*The ramp relative frequency based on a continuous index such as $\{r_t\}$ was calculated by replacing I_t by r_t in equation (3).

can be considered as a robust trade-off between the relative frequencies provided by the reviewed definitions of ramp event (see Figure 2).

The range spanned by $\{R_t\}$ depends on the range spanned by the original wind power time series, which is related to the size of the wind farm. Hence, it is not possible to compare the ramp functions of different wind farms. For this reason, it is necessary to re-scale the ramp function between $[-1, 1]$ by defining the relative ramp function as follows:

$$r_t = \frac{R_t}{\max\{|R_t|\}} \tag{7}$$

In addition, the relative ramp function can be decomposed into three time series, $\{r_t^u\}$, $\{r_t^d\}$ and $\{r_t^0\}$, to isolate the ramp performance during ramp-up (positive ramps), ramp-down (negative ramps) and non-ramp events, respectively:

$$r_t^u = \begin{cases} r_t, & \text{if} \quad r_t \geq 0 \\ 0, & \text{if} \quad r_t < 0 \end{cases} \tag{8}$$

$$r_t^d = \begin{cases} -r_t, & \text{if} \quad r_t \leq 0 \\ 0, & \text{if} \quad r_t > 0 \end{cases} \tag{9}$$

$$r_t^0 = 1 - r_t^u - r_t^d \tag{10}$$

To conclude this subsection, the ramp performance obtained through the ramp function is compared with ramp performances obtained with some of the binary classifications presented in Section 2. Figure 5 shows the wind power time series of Ariz during the period mentioned in Figure 1 (bottom) along with the ramp function computed for $\lambda_N = 5$ (solid line) and $\lambda_N = 10$ (dashed line). In addition, the indicator function for several definitions related to similar sized projects (see Table I) are also represented. Specifically, those time steps labelled as ramp events that would change their status if ΔP_0 is increased by $5\%P_R$ are highlighted with circles. It can be seen that the ramp function takes its minimum values around two ramp-down events, which are also captured by most of the binary criteria. However, disagreements between the binary criteria are observed during strong gradients that would not likely be considered part of a large ramp event. These disagreements are probably because each binary definition was conceived for a specific case study different from the one considered here. However, what is relevant from this figure is that a more reliable ramp characterisation is obtained through a continuous index compared with the ramp characterisations provided by any of the binary classifications: the information provided by the ramp function is not highly sensitive to the λ_N value and provides a means to compare the ramp intensity of two different events without forcing a binary decision on their ramp/non-ramp nature. Conversely, what can be considered a ramp event from a binary point of view may become highly conditioned to the precise threshold value selected, as shown by the circles in the figure. Thus, the ramp function can be considered a satisfactory trade-off between the discrepancies observed under different binary criteria. The ramp function can also provide an alternative to avoiding the sensitivity inherent in the binary classification.

3.3. Discussion on the effect of the addition across scales

In this subsection, we analyse the impact of the proposed addition across scales on the ramp performance by comparing the ramp function defined in equation (6) with respect to other continuous functions, namely the gradient function and the filtered function.

The gradient function of order a, $\{g_t^a\}$, of a certain wind power time series $\{y_t\}$ can be defined as the variation of the wind power generation observed within $a + 1$ time steps:

$$g_t^a = y_t - y_{t-a} \quad , \quad \text{for } t = a + 1, \ldots, N \tag{11}$$

Under the assumption of $\{y_t\}$ being the realisation of a weakly stationary stochastic process $\{Y_t\}$* (commonly adopted in time series analysis,[48]) the expectation of the gradient function is $\mathrm{E}\left[G_t^a\right] = \mathrm{E}\left[Y_t\right] - \mathrm{E}\left[Y_{t-a}\right] = 0$. Let us define the variance of the gradient function as $\mathrm{Var}\left[G_t^a\right] = \sigma_{g,a}^2$. As mentioned previously, most ramp binary definitions are based on the gradient function to define ramp events by setting specific thresholds.[7,11,14,15,19,20,24] However, a continuous ramp performance could be obtained directly by considering $\{g_t^a\}$, since the extreme values of this signal (i.e. the tails of its statistical distribution) would correspond to high variations in power experienced within $a + 1$ time steps. In this case, the value of the parameter a is crucial, since it determines a single time window to be considered. Hence, the use of the gradient function to measure the ramp intensity would be suitable only under the assumption of having a clear idea about a single duration of what can be considered a problematic event.

*Thus, $\mathrm{E}\left[Y_{t1}\right] = \mathrm{E}\left[Y_{t2}\right]$, $\mathrm{Var}\left[Y_{t1}\right] = \mathrm{Var}\left[Y_{t2}\right]$ and $\mathrm{Cov}\left[Y_{t1}, Y_{t1\pm d}\right] = \mathrm{Cov}\left[Y_{t2}, Y_{t2\pm d}\right]$ for every $t1, t2, d$.

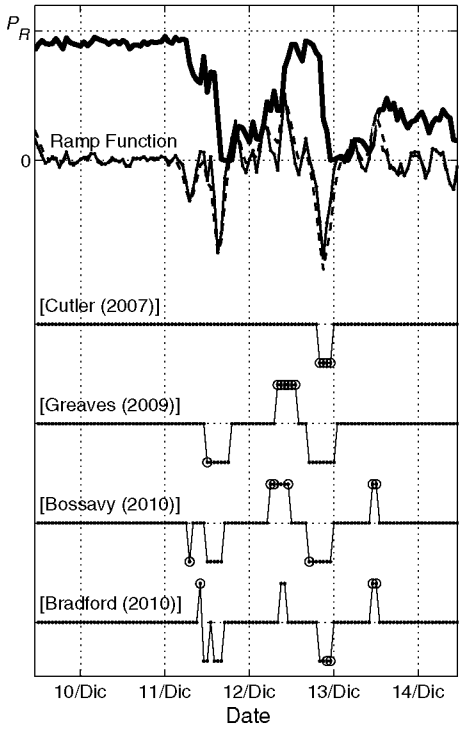

Figure 5. Wind power time series of Ariz with the relative ramp function (solid line for $\lambda_N = 5$ and dashed line for $\lambda_N = 10$) and four indicator functions based on different definitions of a ramp event. Periods labelled as ramp events ($I_t = \pm 1$) that would change to non-ramp ($I_t = 0$) by considering an amplitude threshold S_0 increased by $5\%P_R$ are highlighted with a circle.

A more reliable proposal to characterise the ramp performance of a time series is the filtered function, which averages several gradient functions at each time step. The number and order of the gradient functions considered depend on a scale parameter, λ. The filtered function can be defined as follows:

$$f_t^{\lambda} = \begin{cases} \frac{1}{\sqrt{\lambda}} \cdot \sum_{j=1}^{\lambda/2} g_{t+j-1}^{\lambda/2} & , \text{ if } \lambda \text{ is even} \\ \frac{1}{\sqrt{\lambda}} \cdot \sum_{j=1}^{(\lambda-1)/2} g_{t+j}^{(\lambda+1)/2} & , \text{ if } \lambda \text{ is odd} \end{cases} \tag{12}$$

This definition is equivalent to the filtered signal proposed in Bossavy *et al.*,[21] $\{p_t^f\}$.* In addition, it is worth noting that $\{f_t^{\lambda}\}$ corresponds to the coefficients of the Haar WT for a single scale λ:

$$f_t^{\lambda} = W_H^{t,\lambda}, \ \forall t \tag{13}$$

Since $\{f_t^{\lambda}\}$ is built from the addition of several gradient functions, it can be readily shown that its expectation is $\mathrm{E}\left[F_t^{\lambda}\right] = 0$. Moreover, it can be demonstrated that its variance can be expressed as a linear combination of the variances of several gradient functions as follows:

$$\mathrm{Var}\left[F_t^{\lambda}\right] = \sum_a wf_a \cdot \sigma_{g,a}^2 \tag{14}$$

*$\{p_t^f\}$ was only defined for even scales. The relation among both functions is a lag of one time step and a scalar factor $\sqrt{\lambda}$, that is: $p_t^f(2\lambda) = \sqrt{\lambda} \cdot f_{t+1}^{\lambda}$.

where wf_a is the weight associated to the contribution of the variance of the gradient function $\{g_t^a\}$ to the variance of the filtered function $\{f_t^\lambda\}$. Equations (15) and (16) particularise equation (14) for the case of λ being even and odd, respectively (the demonstration is detailed in Appendix A).

$$Var\left[F_t^\lambda\right] = \frac{1}{2}\cdot\sigma^2_{g,\frac{\lambda}{2}} + \frac{1}{\lambda}\cdot\sum_{j=1}^{\frac{\lambda}{2}-1}\sum_{k=j+1}^{\frac{\lambda}{2}}\left(\sigma^2_{g,\frac{\lambda}{2}+(k-j)} + \sigma^2_{g,\frac{\lambda}{2}-(k-j)} - 2\cdot\sigma^2_{g,(k-j)}\right) \tag{15}$$

$$Var\left[F_t^\lambda\right] = \frac{\lambda-1}{2\lambda}\cdot\sigma^2_{g,\frac{\lambda+1}{2}} + \frac{1}{\lambda}\cdot\sum_{j=1}^{\frac{\lambda-1}{2}-1}\sum_{k=j+1}^{\frac{\lambda-1}{2}}\left(\sigma^2_{g,\frac{\lambda+1}{2}+(k-j)} + \sigma^2_{g,\frac{\lambda+1}{2}-(k-j)} - 2\cdot\sigma^2_{g,(k-j)}\right) \tag{16}$$

Table III (top) gathers the weights wf_a obtained through equations (15) and (16) for several values of λ. It is worth noting that these weights uniquely depend on the way in which the filtered function and the gradient function have been defined (equations (12) and (11), respectively). Hence, the values obtained are valid for any time series $\{y_t\}$ considered.

The reason why the filtered function, $\{f_t^\lambda\}$, provides a more reliable ramp performance than the gradient function, $\{g_t^a\}$, is that $\{f_t^\lambda\}$ is able to communicate high variations in power experienced in several time scales (in particular, in those time scales related to high weights wf_a). In addition, negative weights are related to short time scales where high variations in power would be considered fluctuations rather than ramp events. In this regard, the parameter λ permits to customise to some extent the notion of ramp event by selecting the value that better fits with those time scales where important variations in power would be relevant for a certain study case.

To illustrate the impact of λ on the ramp characterisation obtained with the filtered signal, Figure 6 shows the performance of two filtered signals, $\{f_t^{\lambda=2}\}$ and $\{f_t^{\lambda=6}\}$, during three different events experienced at Ariz. Event (A) shows a ramp event where the power level increased in almost 90% P_R within 3 h. Event (B) illustrates a case of dip in power, i.e. a sharp decrease (B^-) followed by a sharp increase (B^+) in power. Event (C) shows a decrease in power of 75% P_R within 6 h. On the right part of the figure, the histograms of both filtered signals are showed, where the values obtained during the mentioned events have been pointed out. It can be observed that $\{f_t^2\}$ departs from the mean (zero) to extreme values (the tails of its distribution) specially during the events (B^-), (B^+) and (A). On the other hand, $\{f_t^6\}$ highlights events (A) and (C) because they represent high gradients in power evaluated in the range of 4–6 time steps. The ramp characterisation of both filtered functions could be envisaged by visiting the weights wf_a gathered in Table III (top).

Table III. Weights to express the variance of the filtered function, $\{f_t^\lambda\}$, (top) and the ramp function, $\{R_t^{\lambda_N}\}$, (bottom) as a linear combination of the variances of several gradient functions (see equations (14) and (17)).

λ	wf_1	wf_2	wf_3	wf_4	wf_5	wf_6	wf_7	wf_8	wf_9
2	**0.50**	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
3	0.00	**0.33**	0.00	0.00	0.00	0.00	0.00	0.00	0.00
4	−0.25	**0.50**	0.25	0.00	0.00	0.00	0.00	0.00	0.00
5	−0.40	0.20	**0.40**	0.20	0.00	0.00	0.00	0.00	0.00
6	−0.50	0.00	**0.50**	0.33	0.17	0.00	0.00	0.00	0.00
7	−0.57	−0.14	0.29	**0.43**	0.29	0.14	0.00	0.00	0.00
8	−0.62	−0.25	0.12	**0.50**	0.37	0.25	0.12	0.00	0.00
9	−0.67	−0.33	0.00	0.33	**0.44**	0.33	0.22	0.11	0.00
10	−0.70	−0.40	−0.10	0.20	**0.50**	0.40	0.30	0.20	0.10
λ_N	wr_1	wr_2	wr_3	wr_4	wr_5	wr_6	wr_7	wr_8	wr_9
2	**0.50**	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
3	0.50	**0.74**	0.00	0.00	0.00	0.00	0.00	0.00	0.00
4	−0.04	**2.53**	0.54	0.00	0.00	0.00	0.00	0.00	0.00
5	−1.94	**4.01**	2.44	0.42	0.00	0.00	0.00	0.00	0.00
6	−5.08	5.71	**5.98**	1.51	0.35	0.00	0.00	0.00	0.00
7	−9.21	5.41	**9.15**	4.98	1.30	0.30	0.00	0.00	0.00
8	−15.62	4.62	**13.34**	10.00	3.27	1.09	0.26	0.00	0.00
9	−22.64	2.28	14.73	**15.25**	8.18	2.86	0.98	0.23	0.00
10	−31.98	−1.58	16.43	**21.65**	14.89	5.60	2.51	0.86	0.21

The highest value in each row is given in bold fonts.

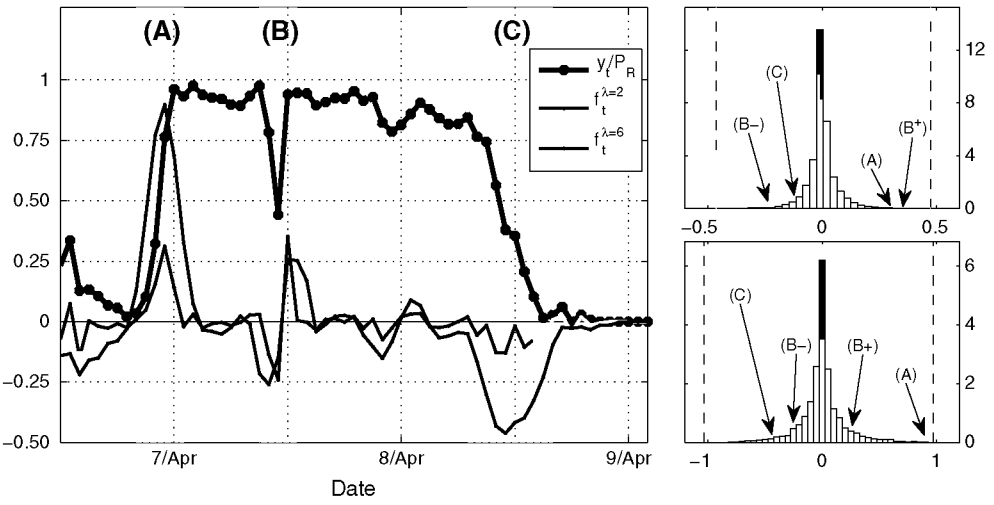

Figure 6. $f_t^{\lambda=2}$ and $f_t^{\lambda=6}$ during three events of high wind power variations (left). Histograms of $f_t^{\lambda=2}$ (right, top) and $f_t^{\lambda=6}$ (right, bottom) with the mentioned events pointed out.

Finally, the case of the ramp function is considered. Taking into account equation (13) along with equation (6), it can be seen that the ramp function is built from the addition of a set of filtered functions with scales λ comprised between λ_1 and λ_N. At this point, the question is to explore the effect of the proposed addition across scales on the characterisation of the ramp performance of a certain time series $\{y_t\}$. From now on, we will consider the case of $\lambda_1 = 2$ since it represents the shortest scale to evaluate a gradient (two time steps). To simplify the notation, let us denote the ramp function $R_t(\lambda_1 = 2, \lambda_N)$ with $\{R_t^{\lambda_N}\}$. It can be noted from equations (6) and (13) that the ramp function can be expressed as a linear combination of several gradient functions. Likewise, the variance of the ramp function can be expressed as a linear combination of the variances of several gradient functions, $\{g_t^a\}$, as follows:

$$Var\left[R_t^{\lambda_N}\right] = \sum_a wr_a \cdot \sigma_{g,a}^2 \tag{17}$$

where wr_a is the weight associated to the contribution of the variance of the gradient function $\{g_t^a\}$ to the variance of the ramp function. A general expression for these coefficients becomes excessively complex in this case. Nevertheless, they can be computed through equations (6), (12), (13), (22) and (28). Table III (bottom) gathers the weights wr_a for several values of λ_N.

The value of λ_N determines (through the weights wr_a) those time scales where large variations in power would be highlighted by $\{R_t^{\lambda_N}\}$, providing a parameter to customise the notion of ramp event for a specific study case. However, the ramp characterisation provided by the ramp function differs from the one provided by the filtered function on the fact that wr_a concentrates within a narrow range of time scales, whereas wf_a evolves more regularly along the scales a. This effect can be observed in Figure 7, which illustrates the weights obtained wf_a and wr_a for several cases of λ and λ_N (to facilitate the comparison, each set of weights was divided by its maximum value). For example, let consider the case of $\lambda = \lambda_N = 10$. By visiting the coefficients wr_a in Table III, it can be noted that extreme values of the ramp function, $\{R_t^{10}\}$, are likely to be related to extreme values of the gradient functions $\{g_t^3\}$,$\{g_t^4\}$ and $\{g_t^5\}$. However, large gradients experienced within other time scales would have little impact on the ramp function because of the lower value of the related weighs. On the other hand, extreme values of the filtered function, $\{f_t^{10}\}$, are likely to be related to high variations in power experienced under a broader number of time scales because of the smooth decrease of the weights wf_a with respect to the scale a. Therefore, it can be concluded that both the filtered function and the ramp function permit to highlight large variations in power experienced within a certain range of time scales, providing a better ramp characterisation than the gradient function. However, the addition across scales proposed in the ramp function definition contributes in a better bound of the range of time scales involved in the notion of ramp event.

4. PROPOSALS FOR THE USE OF THE RAMP FUNCTION

The ramp function defined previously is intended to characterise the ramp performance of a wind power time series in a more reliable way than binary classification because it is less sensitive to input parameters and a continuous index related

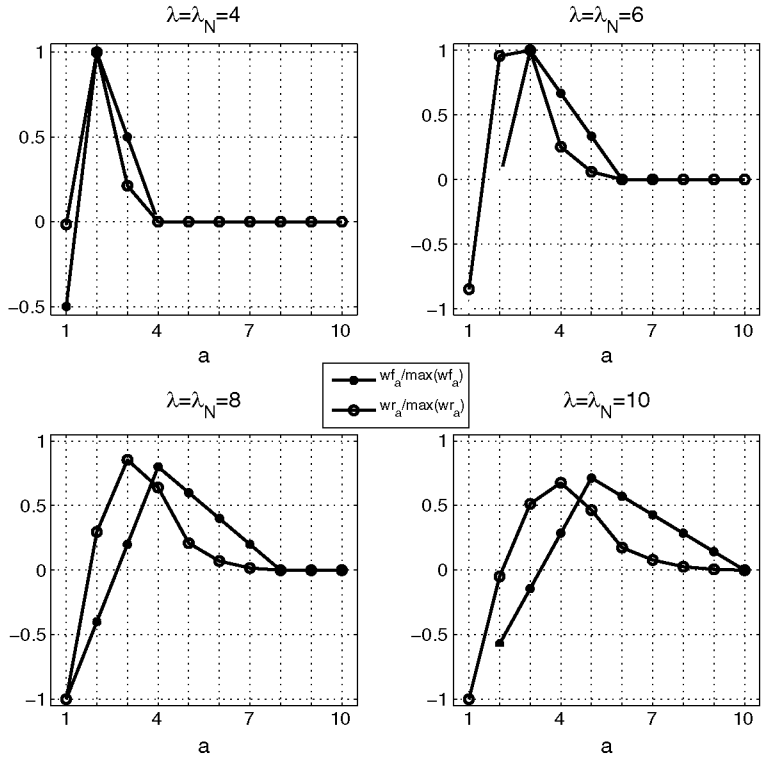

Figure 7. Weights wf_a and wr_a (defined in equations (14) and (17) for several cases of λ and λ_N. To facilitate the comparison, each set of weights was divided by its maximum value.

to the ramp intensity is provided at each time step. Both end-users and forecaster/modellers could take advantage of the ramp function to enhance their approach to wind power ramp forecasting, although specific considerations must be taken into account. For example, concerning the labour of end-users, the ramp function may represent a suitable option for better communicating the impact of ramp events to forecaster/modellers because it would be possible to tailor the impact level to different ramp intensities. However, this option would require the end-users to explore the interpretation of the ramp function indexes in terms of their specific cost function. From the forecaster/modellers point of view, the ramp function provides a characterisation of the ramp performance mainly based on the considered wind power time series, whereas the information obtained through a binary classification is subjected to the threshold values. Furthermore, the continuous indexes provided by the ramp function permit new approaches to wind power ramp analysis that are not possible from a binary classification. Nevertheless, it is important to note that the evaluation of $\{R_t\}$ at a certain time $t = t_0$ requires a set of λ_N records of $\{y_t\}$ centred in $\{y_{t_0}\}$. Thus, $\{R_{t_0}\}$ would be available $\lambda_N/2$ time steps after $t = t_0$. This fact does not represent a problem for preliminary offline analyses of ramp causes identification, which can be used to customise ramp forecasting tools when considering a certain location. However, this delay has to be taken into account for the case of online applications (e.g. computing the ramp function of the forecasted wind power time series to provide online alerts).

The aim of the following subsections is to propose several applications based on the ramp function that concern the offline analysis of ramp events observed in wind power time series. We therefore focus on the forecaster/modellers perspectives. End-users' motivations are not explored further in this article, although this perspective represents an interesting path for research in future studies.

4.1. Characterising the diurnal pattern of ramp events

Atmospheric dynamics usually shows a component of diurnal pattern due to solar radiation forcing. Thus, it would be of interest to explore whether this pattern is also present in ramp events. This analysis could eventually give insights into the nature of certain ramp events at a given site.

Let us consider a wind power time series with the length of 1 year length from the Ariz wind farm, $\{y_t\}$, along with the relative ramp functions $\{r_t^u\}$, $\{r_t^d\}$ and $\{r_t^0\}$ defined in equations (8–10). Let us define $\{r_*^u\}$ as the set of 8760 elements

contained in $\{r_t^u\}$ sorted in decreasing order. Similar definitions correspond to $\{r_*^d\}$ and $\{r_*^0\}$. Given that $\{r_t^u\}$ and $\{r_t^d\}$ contain the values provided by the relative ramp function during ramp-up and ramp-down events, respectively, $\{r_*^u\}$ and $\{r_*^d\}$ allow us to compare differences (in terms of frequency and intensity) among the two possible directions of ramp events. Figure 8 (top) shows $\{r_*^u\}$, $\{r_*^d\}$ and $\{r_*^0\}$ overlapped and illustrates that for most of the year, the value provided by the relative ramp functions is negligible. In addition, Figure 8 (bottom) reveals in detail that ramp-up events were more intense than ramp-down events. A similar conclusion was met in Katzenstein *et al.*[49] for an interconnected area of 20 wind farms in Texas.

Now, let us define $\{r_{t:h}^u\}$ as the set of values that result from considering only those values of $\{r_t^u\}$ occur at time h of each day, given in the following equation:

$$r_{t:h}^u = \{r_t^u | t = h + 24 * i\}, \quad \text{for} \quad i = 1, \ldots, 365 \tag{18}$$

Once more, each $\{r_{t:h}^u\}$ can be sorted in decreasing order to generate the corresponding $\{r_{*:h}^u\}$. By performing this sort, the first element of $\{r_{*:10}^u\}$ provides the maximum value of the ramp function obtained at 10:00 h, and so forth. Figure 9 illustrates the 24 time series $\{r_{*:h}^u\}$ (green) along with the 24 time series $\{r_{*:h}^d\}$ (red). Thus, the observed colour determines whether the observed ramp-up or the ramp-down events were more intense (higher ramp function values) depending on the hour of the day considered. In our case, it can be seen that ramp-down events were more intense for the part of the day comprised from 08:00 to 16:00 h, whereas the opposite situation was observed for the rest of the hours. For more clarity, Figure 10 shows four sections of Figure 9. This result suggests that ramp events at this wind farm follow a diurnal pattern. Hence, for this specific case study, daily cycles are likely to play a key role in explaining ramp events' underlying causes. Finally, further research can be carried out to explore different cycles of interest, such as seasonal cycles.

4.2. Analysis of a wind farm's ramp performance at the wind turbine level

Since the wind power output of a certain wind farm is generated from the addition of the power delivered by a number of wind turbines, the wind power output dynamics of each wind turbine influence to some degree those dynamics observed in the wind power output of the wind farm. However, considering that each wind turbine is affected by local conditions (e.g. the local orography and the relative position with respect to the other wind turbines), the contribution of each individual

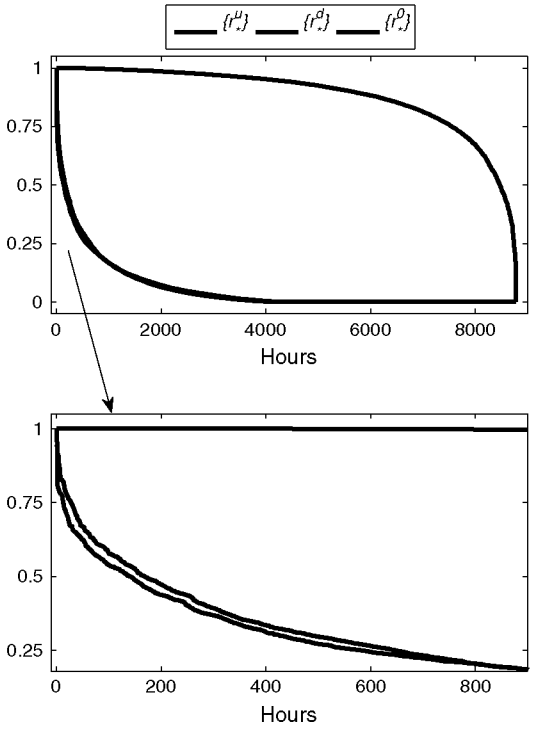

Figure 8. Relative ramp functions sorted in descending order (top). Observed ramp-up events were more intense than ramp-down events (bottom).

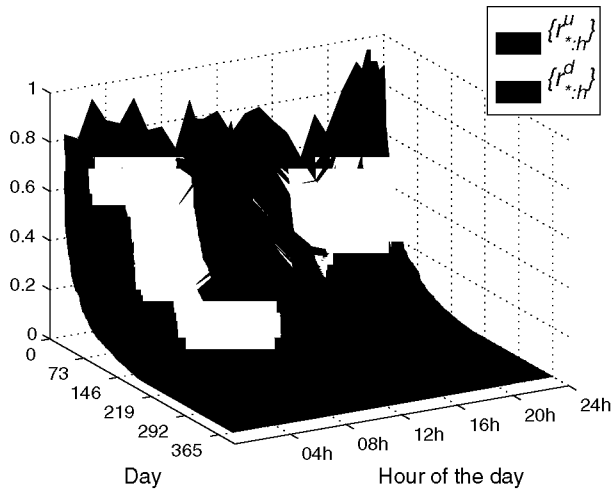

Figure 9. $\{r^u_{*;h}\}$ and $\{r^d_{*;h}\}$. The colour observed for each hour of the day reveals whether ramp-up or ramp-down events were more intense.

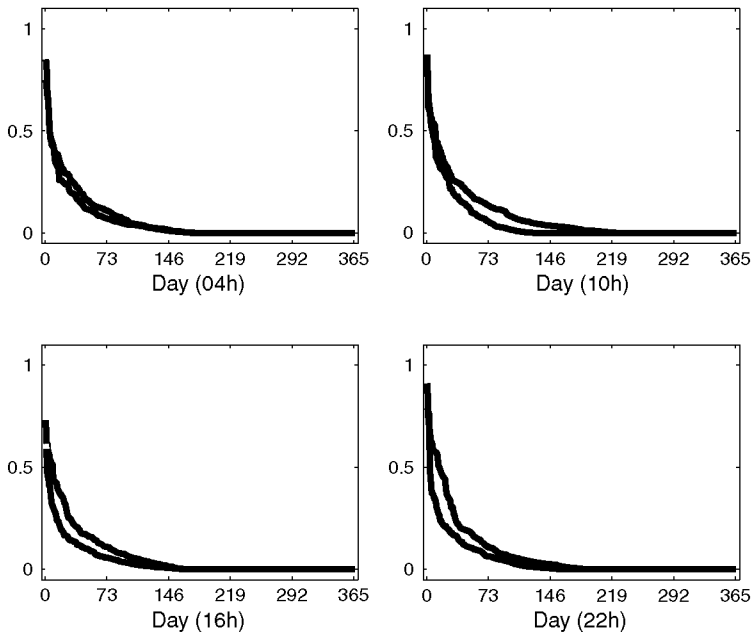

Figure 10. Four sections of Figure 9.

wind turbine to the dynamics of the total wind power output may vary from one turbine to another. In this regard, the ramp function allows us to evaluate the correlation of the ramp performance among different wind turbines within a wind farm.

For this experiment, the case study of a wind farm located in La Muela (Spain) was considered. The wind farm consists of 65 wind turbines. The hourly wind power time series is available for each wind turbine for 1 year. Figure 11 shows the lay out of the wind turbines and the wind rose. A previous study of this wind farm showed that several groups of wind turbines (clusters) can be identified depending on different power performances.[50] Similarly, the use of the ramp function may eventually give insights about the different levels of ramp performance. Let us consider the ramp functions $\{R^i_t\}$, $i = 1, \ldots, 65$ defined in equation (6) for each wind turbine power output time series. The cross-correlation function $C_{ij}(l)$ can be readily obtained from equation (19) for each pair of wind turbines to measure the similarity of their ramp behaviour in a more accurate way compared with correlating their wind power time series, since the ramp function departs from zero depending on the intensity of the ramp event experienced by a certain wind turbine. As a result, high correlation

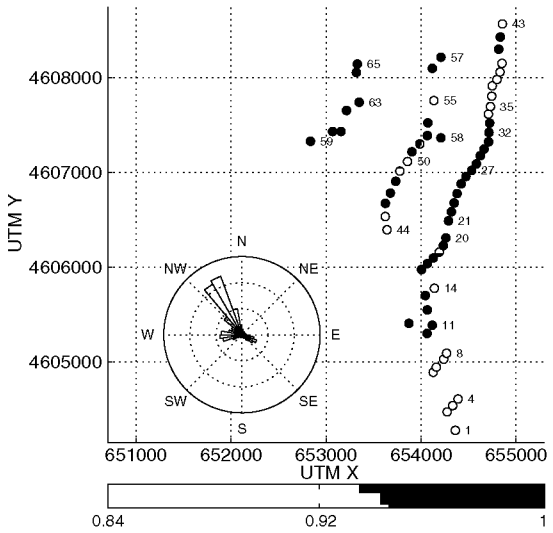

Figure 11. Layout of the La Muela wind farm. The colour represents the cross-correlation between the ramp function of each wind turbine and the ramp function of the wind farm power output.

coefficients would inform us about ramp events taking place at different wind turbines with a given lag l.

$$C_{ij}(l) = \frac{1}{(N-l-1)\cdot\sigma_{R^i}\cdot\sigma_{R^j}} \sum_{t=1+l}^{N} (R_t^i - \overline{R^i})\cdot(R_{t-l}^j - \overline{R^j}) \tag{19}$$

In equation (19), N is the length of the time series, $\overline{R^i}$ is the sample mean of $\{R_t^i\}$, σ_{R^i} is the sample standard deviation of R_t^i and l is the lag considered between both time series. The maximum of $C_{ij}(l)$ was found to be at $l = 0$ for all pair (i, j), which is consistent with the fact that ramp event causes (generally based on meteorological effects) take less than 1 h to propagate across the whole wind farm (although a different conclusion could be obtained for larger wind farms or portfolios).* Figure 12 illustrates $C_{ij}(0)$, which ranges from a minimum of 0.68 to 1. It is possible to observe clusters of wind turbines with highly correlated ramp functions (wind turbines from 21 to 25), which correspond to the central part of the third wind turbine line. In contrast, the ramp performance of wind turbines 1 to 5 (located at the extreme edge of the first line) has a low correlation compared with wind turbines 34 to 43 (at the opposite edge with respect to the prevailing wind direction).

The cross-correlation function between the ramp function of each wind turbine and the ramp function of the wind farm was also estimated (only the case $l = 0$ was considered). The correlations ranged from 0.85 to 0.96, showing the highest values for the wind turbines located at the centre of the third row (wind turbines 21–27; see Figure 11). This fact suggests that there is a significant connection between ramp events observed in the total power output time series and ramp events experienced in a small part of the wind farm. Hence, special attention could be paid to this cluster during ramp events so as to provide insights about the ramp causes. Specifically, it would be of interest to explore the contribution of microscale effects (such as wakes) that affect this cluster to the generation of wind power ramp events at this wind farm. Another interesting issue would be to evaluate the contribution of online measurements of this cluster to the performance of ramp forecasting models rather than the power output of the whole wind farm.

4.3. Assessing the performance of statistical point-forecasting models during ramp events

The aim of statistical point-forecasting models is to learn and replicate from a statistical perspective the dynamics shown by a certain time series. This kind of models usually overcome Numerical Weather Prediction models for very short-term wind power forecasting (from a few minutes to a few hours.[4,5]) There are several criteria for evaluating the performance

*An estimation of the characteristic time T_{char} of a wind perturbation crossing the wind farm is $T_{\text{char}} \simeq L_{\text{char}}/V_{\text{ave}}$, where L_{char} is a characteristic length of the wind farm and V_{ave} is the average wind speed. In this case, $T_{\text{char}} << 1$ h.

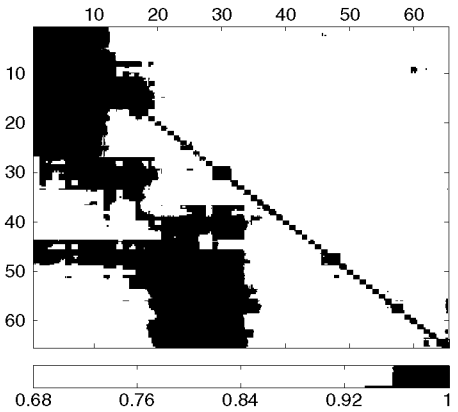

Figure 12. Cross-correlations between ramp functions of the 64 wind turbines in the La Muela wind farm.

of wind power statistical point-forecasting models.[51] One of the most often employed criterion is the mean squared error (*MSE*), which is defined as follows:

$$MSE(k) = \frac{1}{N-k} \sum_{t=1}^{N-k} (y_{t+k} - \hat{y}_{t+k|t})^2 \tag{20}$$

where k is the prediction horizon, N is the number of observations of $\{y_t\}$ and $\hat{y}_{t+k|t}$ is the forecast of y_{t+k} provided by a statistical model at time t. The *MSE* is a global criterion, which means that the performance of the model is evaluated across a complete period (1 year, for instance). However, it may be of interest to evaluate different models analysing their performance during specific events such as ramp events. This task can be achieved by employing the relative ramp functions $\{r_t^u\}$, $\{r_t^d\}$ and $\{r_t^0\}$, which are defined in equations (8–10). The underlying idea is to weight the forecast error incurred at each time step with these functions. For example, the MSE^u defined in equation (21) would correspond to the *MSE* committed during ramp-up events.

$$MSE^u(k) = \frac{1}{\sum_{t=1}^{N-k} (r_{t+k}^u)} \sum_{t=1}^{N-k} r_{t+k}^u \cdot (y_{t+k} - \hat{y}_{t+k|t})^2 \tag{21}$$

Similar definitions would follow for $MSE^d(k)$ (*MSE* during ramp-down events) and $MSE^0(k)$ (*MSE* during non-ramp events). It is worth noting that this way of weighting takes advantage of the fact that the relative ramp function provides a more accurate set of weights than a binary classification of ramp events, since the latter would provide a set of weights on the basis of zeros and ones whereas the relative ramp function takes different values depending on the intensity of each ramp event. In addition, the binary definition may become highly dependent on the thresholds used to define a ramp, as shown in Section 2.1.

To illustrate the evaluation of the *MSE* during ramp events, the case study of a wind farm located in the complex region of Alaiz (northwest Spain) was considered. The wind power time series with an hourly resolution has been derived from 3 years of available data. The following sets of prediction models have been proposed for one step-ahead forecasting ($k = 1$):

- Persistence, which is the main reference method for short-term forecasting. Persistence states that the forecast is the last observed value. Considering the case of $k = 1$,

$$\hat{y}_{t+1} = y_t$$

- Linear autoregressive AR(p) model. The forecast is a linear combination of the last p observed values, given by

$$\hat{y}_{t+1} = \varphi_0 + \sum_{i=1}^{p} \varphi_i \cdot y_t$$

The coefficients φ_i were estimated by means of the Yule-Walker equation, and the optimal AR order p was found to be $p = 3$.

- Artificial Neural Network model (ANN). ANNs have become a powerful computational model because they allow non-linear patterns to be captured simply from a set of historical data. In this case, a multilayer perceptron (MLP) was used. The appropriate architecture of the MLP and the optimal number of inputs (last observed values) were determined from a cross-validation criterion. As a result, the forecast at time $t+1$ is provided by a non-linear function F as follows:

$$\hat{y}_{t+1} = F(y_t, y_{t-1}, y_{t-2})$$

- Regime-Switching Artificial Neural Nets. This model was proposed in Gallego *et al.*[24] The underlying idea is to combine the power of ANN models with a regime-switching strategy. Three regimes were identified in the wind power time series depending on the last observed gradient $y_t - y_{t-1}$. A different MLP was trained for each regime:

$$\hat{y}_{t+1} = \begin{cases} F_1(y_t, y_{t-1}), & \text{regime 1} \\ F_2(y_t, y_{t-1}, y_{t-2}), & \text{regime 2} \\ F_3(y_t, y_{t-1}, y_{t-2}, y_{t-3}), & \text{regime 3} \end{cases}$$

A period of 2 years was employed to train and validate the models. Next, the *MSE* was evaluated during the third year. Figure 13 shows the *MSE* for each model depicted in ramp-up, ramp-down and non-ramp events. Several conclusions can be derived from this figure. (i) It can be observed that forecast errors are, as might be expected, higher during ramp events than during non-ramp events for every model. (ii) Considering the performance of persistence, the lower *MSE* during ramp-down events compared with the *MSE* obtained during ramp-up events suggests that the ramp intensity is statistically higher during ramp-up events compared with that during ramp-down events. This conclusion is similar to the conclusion in Section 4.1 for the wind farm of Ariz and the case study of Texas considered in Katzenstein *et al.*[49] (iii) With regard to the results obtained by the linear AR, the ANN and the RS-ANN models, it was found that non-linear models allowed important reductions of the *MSE* during ramp-up events. Specifically, the ANN and the RS-ANN obtained a reduction of 5.6% and 9.3% with respect to the AR model. Considering the results for the ramp-down events, it can be noticed that more complex models did not attain similar *MSE* reduction levels with respect to the *MSE* obtained with the linear AR model (2.5% and 5.2%, respectively). The fact that more complex models did not produce better forecasts during ramp-down events does not permit us to conclude whether the linear model performs sufficiently well to capture the wind power dynamics during these events or that hypothetical room for improvement was not exploited by the proposed non-linear models. The second option would be supported by the fact that ramp-down events could be due to a higher diversity of underlying causes, such as wind speeds passing the cut-out speed of wind turbines or yaw misalignments caused by sudden changes in the wind direction. In this regard, a lower predictability during ramp-down events was pointed out in Truewind.[15] (iv) Finally, the accuracy of the models was practically the same during non-ramp events, suggesting that complex models do not seem to provide a substantial contribution with respect to Persistence.

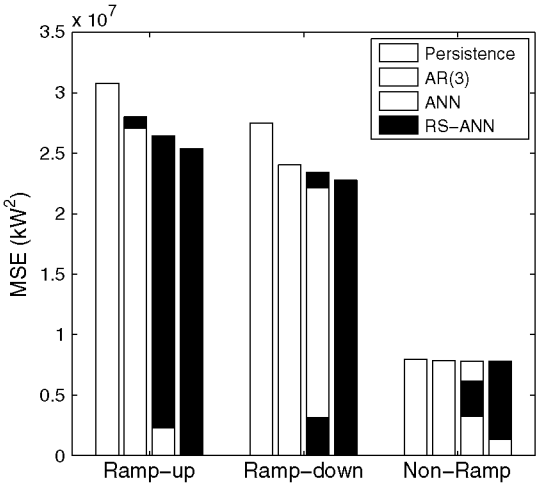

Figure 13. Evaluation of the wind power forecasting performance of different statistical models for three situations: ramp-up events, ramp-down events and non-ramp events.

5. CONCLUSIONS

The present study introduced a novel tool based on the WT. This tool is used to characterise the ramp performance of wind power time series. The underlying concept is to consider the contribution of the local gradient evaluated under different time scales. As a result, the ramp function is obtained, which provides a continuous index related to the ramp intensity at each time step.

The use of a continuous index was found to be a more reliable approach to characterise the ramp performance of wind power time series than the binary classification ramp/non-ramp usually employed in wind power ramp forecasting. Although the ramp function does not provide a 'closed definition' of ramp events, it permits us to address several drawbacks related to the binary classification. For example, the ramp function is based on the assumption that there is an imprecise border that differentiates ramp events within a wind power time series. This assumption is important to minimise the sensitivity observed when binary classification is employed for ramp identification. Hence, specific thresholds regarding fixed values on power variations and ramp durations are not required. Instead, the ramp function can be computed by considering one parameter related to relevant time scales in the ramp characterisation. Moreover, the ramp function is obtained in such a way that small variations in this parameter do not entail substantial variations in the ramp intensity indexes.

On the other hand, binary classification supports the notion that ramp events are similar to one another, although differences in amplitude and duration are usually observed. In contrast, the ramp function preserves the information in these features by assessing a continuous index of the intensity of the ramp experienced at each time step. This fact potentially benefits the activity of both end-users (TSOs and energy traders) and forecaster/modellers in relation to wind power ramp forecasting.

From the point of view of forecasters/modellers, the ramp function permits new approaches to wind power ramp analysis that are not feasible from a binary classification standpoint. To illustrate this fact, a set of applications where the ramp function plays a major role was described. Interesting conclusions concerning different aspects of wind power ramp forecasting were envisaged for some wind farms located in Spain. The first application shows how the proposed ramp function could be used to identify the period of the day in which ramp-up events are more intense compared with ramp-down events and vice versa. The second application explores how a specific cluster of wind turbines could be identified in a wind farm on the basis of how close its ramp performance is to the ramp performance of the whole wind farm. Finally, the third application shows how the ramp function could contribute to evaluate statistical point-forecasting models during ramp events, providing insights into different predictability levels for ramp-up and ramp-down events. Nevertheless, there is still room for the development of new approaches. For example, considering that ramp events may be motivated by a broad number of causes, an important open question is how much the proposed ramp function could eventually contribute to a better understanding of the meteorological processes and the operational states that cause ramp events. Future research could address this topic as the next step for the development of oriented ramp event forecasting tools.

The issue of the ramp function contribution to end-users was not addressed in depth in this work and should be further investigated. However, it can be envisaged that the ramp function would permit end-users to tailor different ramp intensities to a specific cost function. In this regard, energy traders and TSOs could eventually benefit from the ramp function since a more appropriate action could be adopted depending on the forecasted ramp intensity rather than the limited range of information provided by the binary classification. The counterpart would be that the end-user needs to be more knowledgeable about the implications of a particular ramp intensity (in terms of economic penalties or the cost of ancillary services).

Finally, it is important to note that the theoretical background based on the WT permits the ramp function to be readily extended to other time series, such as wind speed or wind direction time series, as well as to time series with higher time resolutions than the hourly resolution used in this work, to address the characterisation of faster ramp events.

ACKNOWLEDGEMENTS

This research was carried out during the AVAVIP Project (CGL2008-05-093/CLI) funded by the Ministry of Science and Innovation in the frame of the Spanish National Plan for Research and Development 2008/2010. The first author would like to express his gratitude to Ciemat who funded part of his PhD programme. Acciona is also acknowledge for providing the wind power measurements of several wind farms employed during this study. Finally, the authors would like to thank to the reviewers for their valuable comments and suggestions during the revision process.

APPENDIX A

In this Appendix, equations (15) and (16) are demonstrated. Consider a time series, $\{y_t\}$, obtained through the realisation of the stochastic process $\{Y_t\}, t = 1, \ldots, N$, which is supposed to be weakly stationary (i.e. $\mathrm{E}[Y_{t1}] = \mathrm{E}[Y_{t2}]$,

$Var[Y_{t1}] = Var[Y_{t2}]$ and $Cov[Y_{t1}, Y_{t1\pm d}] = Cov[Y_{t2}, Y_{t2\pm d}]$ for every $t1, t2, d$). The demonstration makes use of some properties related to the variance and covariance of several random variables, namely

$$Var\left[\sum_{j=1}^{n} X_j\right] = \sum_{j=1}^{n} Var[X_j] + 2 \cdot \sum_{j=1}^{n-1} \sum_{k=j+1}^{n} Cov[X_j, X_k] \tag{22}$$

$$Cov[aX, bY] = a \cdot b \cdot Cov[X, Y] \tag{23}$$

$$Cov[X + Y, Z] = Cov[X, Z] + Cov[Y, Z] \tag{24}$$

$$Cov[X, Y] = Cov[Y, X] \tag{25}$$

The variance of the gradient function defined in equation (11) is given by

$$Var\left[G_t^a\right] = Var[Y_t] + Var[Y_{t-a}] - 2 \cdot Cov[Y_t, Y_{t-a}] = 2 \cdot Var[Y_t] - 2 \cdot Cov[Y_t, Y_{t-a}] \tag{26}$$

Equation (26) can be rewritten as follows:

$$Cov[Y_t, Y_{t-a}] = \frac{2 \cdot Var[Y_t] - Var\left[G_t^a\right]}{2} \tag{27}$$

Now, let consider the covariance between two gradient functions, $\left[G_t^a\right]$ and $\left[G_{t-d}^{a-b}\right]$, for arbitrary values a, b and d:

$$\begin{aligned} Cov\left[G_t^a, G_{t-d}^{a-b}\right] &= Cov\left[Y_t - Y_{t-a}, Y_{t-d} - Y_{t-d-a+b}\right] \\ &= Cov[Y_t, Y_{t-d}] - Cov\left[Y_t, Y_{t-d-a+b}\right] - Cov\left[Y_{t-a}, Y_{t-d}\right] + Cov\left[Y_{t-a}, Y_{t-d-a+b}\right] \\ &\overset{(27)}{=} \frac{Var\left[G_t^{d+a-b}\right] - Var\left[G_t^d\right] + Var\left[G_t^{a-d}\right] - Var\left[G_t^{d-b}\right]}{2} \end{aligned} \tag{28}$$

Finally, consider the filtered signal defined in equation (12) for the case of λ being even (the demonstration for λ being odd is analogous). The variance of the filtered signal can be written as follows:

$$\begin{aligned} Var\left[F_t^{\lambda}\right] &= Var\left[\frac{1}{\sqrt{\lambda}} \cdot \sum_{j=1}^{\lambda/2} G_{t+j-1}^{\lambda/2}\right] \\ &= \frac{1}{\lambda} \cdot \left(\sum_{j=1}^{\lambda/2} Var\left[G_{t+j-1}^{\lambda/2}\right] + 2 \cdot \sum_{j=1}^{\lambda/2-1} \sum_{k=j+1}^{\lambda/2} Cov\left[G_{t+j-1}^{\lambda/2}, G_{t+k-1}^{\lambda/2}\right]\right) \\ &\overset{(28)}{=} \frac{1}{2} \cdot Var\left[G_t^{\lambda/2}\right] + \frac{2}{\lambda} \cdot \sum_{j=1}^{\lambda/2-1} \sum_{k=j+1}^{\lambda/2} \frac{Var\left[G_t^{\lambda/2+k-j}\right] - 2 \cdot Var\left[G_t^{k-j}\right] + Var\left[G_t^{\lambda/2-k+j}\right]}{2} \end{aligned} \tag{29}$$

which leads to equation (15).

REFERENCES

1. IEA. Design and operation of power systems with large amounts of wind power. *Final report, IEA WIND Task 25, Phase one 2006–2008*, 2009. Available (23-Mar-2011) at http://www.ieawind.org/AnnexXXV/PDF/Final%20Report%20Task%2025%202008/T2493.pdf.
2. Landberg L, Giebel G, Nielsen H, Nielsen T, Madsen H. Short-term prediction—an overview. *Wind Energy* 2003; **6**(3): 273–280. DOI: 10.1002/we.96.
3. Smith JC, Milligan MR, DeMeo EA, Parsons B. Utility wind integration and operating impact state of the art. *IEEE Transactions on Power Systems* 2007; **22**(3): 900–908. DOI: 10.1109/TPWRS.2007.901598.

4. Costa A, Crespo A, Navarro J, Lizcano G, Madsen H, Feitosa E. A review on the young history of the wind power short-term prediction. *Renewable and Sustainable Energy Reviews* 2008; **12**(6): 1725–1744. DOI: 10.1016/j.rser.2007.01.015.
5. Giebel G. The state of the art in short-term prediction of wind power—a literature overview (2nd ed.) *Technical Report*, ANEMOS.plus/SafeWind projects, 2011. Available at http://www.prediktor.dk/.
6. Hawkins D, Rothleder M. Evolving role of wind forecasting in market operation at the CAISO, *IEEE/PES Power Systems Conference and Exposition*, Atlanta, GA, 2006, DOI: 10.1109/PSCE.2006.296304, (to appear in print).
7. Potter CW, Grimit E, Nijssen B. Potential benefits of a dedicated probabilistic rapid ramp event forecast tool, *IEEE/PES Power Systems Conference and Exposition*, Seattle, WA, 2009, DOI: 10.1109/PSCE.2009.4840109, (to appear in print).
8. Giebel G. The new state-of-the-art report and the European experience in safe wind and other projects, *4th Workshop on Best Practice in the Use of Short-term Forecasting of Wind Power*, Quebec, 2010. Available (23-Mar-2011) at http://powwow.risoe.dk/BestPracticeWorkshop4.htm.
9. 2011. Available at http://www.safewind.eu/ (Accessed 23 March 2011).
10. Thiesen J. Nowcasting ramp events, *Third Workshop on Best Practice in the Use of Short-term Forecasting of Wind Power*, Bremen, 2009. Available (23-Mar-2011) at http://powwow.risoe.dk/publ/JThiesen_(ConWx)-NowcastingRampEvents_BestPracticeSTP-3_2009.pdf.
11. Collier C, Parkes J, Collins J, Landberg L. Improved ramp event forecasting using upstream wind measurements, *European Wind Energy Conference*, Warsaw, 2010.
12. Zack JW, Aymamí J, Vidal-Pérez J, Tortosa-Andreu A. *An Innovative Short-Term Large Wind Ramp Forecasting System*, Geophysical Research Abstracts, Vol. 12, EGU General Assembly, 2010.
13. Girard R, Bossavy A, Kariniotakis G. Forecasting ramps of wind power production at different time scales, *European Wind Energy Conference*, Brussels, 2011.
14. Cutler N, Kay M, Jacka K, Nielsen TS. Detecting, categorizing and forecasting large ramps in wind farm power output using meteorological observations and WPPT. *Wind Energy* 2007; **10**(5): 453–470. DOI: 10.1002/we.235.
15. Truewind A. AWS Truewind's final report for the Alberta forecasting pilot project, 2008. Available (23-Mar-2011) at http://www.uwig.org/Alberta_PP_Final_Report_AWST_Jun25.pdf.
16. Zack JW. Optimization of wind power production forecast performance during critical periods for grid management, *AWEA Windpower Conference*, Los Angeles, CA, USA, 2007. Available (23-Mar-2011) at http://www.awstruepower.com/wp-content/media/2010/05/AWEA_Windpower_2007_Forecasting.pdf.pdf.
17. Cutler NJ, Outhred HR, MacGill IF, Kay MJ, Kepert JD. Characterizing future large, rapid changes in aggregated wind power using Numerical Weather Prediction spatial fields. *Wind Energy* 2009; **12**(6): 542–555. DOI: 10.1002/we.312.
18. Martí I. Introduction to meteorology: forecasting, *Talk at the First European Academy of Wind Energy - Wind Resource Assessment Audit Standardization (EAWE-WAUDIT) project Summer School*, Pamplona, 2010.
19. Greaves B, Collins J, Parkes J, Tindal A. Temporal forecast uncertainty for ramp events. *Wind Engineering* 2009; **33**(4): 309–320. DOI: 10.1260/030952409789685681.
20. Bradford K, Carpenter R, Shaw B. Forecasting southern plains wind ramp events using the WRF model at 3-Km, *The 9th American Meteorological Society Annual Meeting*, Atlanta, GA, 2010. Available (23-Mar-2011) at www.caps.ou.edu/reu/reu09/papers/Bradford.pdf.
21. Bossavy A, Girard R, Kariniotakis G. Forecasting wind power uncertainty related to ramp events, *European Wind Energy Conference*, Warsaw, 2010.
22. Pease J. Critical short-term forecasting needs for large and unscheduled wind energy on the BPA system, *Conference at the 3rd Workshop on Best Practice in the Use of Short-term Forecasting of Wind Power*, Bremen, 2009. Available (23-Mar-2011) at http://powwow.risoe.dk/publ/JPease_(BPA)-BPAWindRampEventTrackingSystem_BestPracticeSTP-3_2009.pdf.
23. Hedley TB, Keffer JaF. Turbulent/non-turbulent decisions in an intermittent flow. *Journal of Fluid Mechanics* 1974; **64**(04): 625–644. DOI: 10.1017/S0022112074001832.
24. Gallego C, Costa A, Cuerva A. Improving short-term forecasting during ramp events by means of Regime-Switching Artificial Neural Networks. *Advances in Science and Research* 2011; **6**: 55–58. DOI: 10.5194/asr-6-55-2011.
25. Pinson P. Catalogue of complex to extreme situations. SafeWind project Deliverable, 2009. Available at http://www.safewind.eu/images/Articles/Deliverables/.
26. Daubechies I. *Ten lectures on wavelets*, (1st edn), SIAM: Society for Industrial and Applied Mathematics, 1992.

27. Morlet J, Arens G, Fourgeau E, Glard D. Wave propagation and sampling theory. *Geophysics* 1982; **47**(2): 203–236. DOI: 10.1190/1.1441328.
28. Akansu AN, Serdijn WA, Selesnick IW. Emerging applications of wavelets: a review. *Physical Communication* 2010; **3**(1): 1–18. DOI: 10.1016/j.phycom.2009.07.001.
29. Amolins K, Zhang Y, Dare P. Wavelet based image fusion techniques—an introduction, review and comparison. *ISPRS Journal of Photogrammetry and Remote Sensing* 2007; **62**(4): 249–263. DOI: 10.1016/j.isprsjprs.2007.05.009.
30. Canny J. A computational approach to edge detection. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 1986; **8**: 679–698. DOI: 10.1109/TPAMI.1986.4767851.
31. Grossmann A. Wavelet transforms and edge detection. In *Stochastic processes in physics and engineering*, et al. SA (ed.). D. Reidel Publishing Company: Dordrecht, 1988; 149–157.
32. Mallat S, Hwang WL. Singularity detection and processing with wavelets. *IEEE Transactions on Information Theory* 1992; **38**(2): 617–643.
33. Milidiú RL, Machado RJ, Rentería RP. Time-series forecasting through wavelets transformation and a mixture of expert models. *Neurocomputing* 1999; **28**: 145–156. DOI: 10.1016/S0925-2312(98)00120-9.
34. Yao SJ, Song YH, Zhang LZ, Cheng XY. Wavelet transform and neural networks for short-term electrical load forecasting. *Energy Conversion and Management* 2000; **41**(18): 1975–1988. DOI: 10.1016/S0196-8904(00)00035-2.
35. Zhang B, Dong Z. An adaptive neural-wavelet model for short term load forecasting. *Electric Power Systems Research* 2001; **59**(2): 121–129. DOI: 10.1016/S0378-7796(01)00138-9.
36. Tan Z, Zhang J, Wang J, Xu J. Day-ahead electricity price forecasting using wavelet transform combined with ARIMA and GARCH models. *Applied Energy* 2010; **87**(11): 3606–3610. DOI: 10.1016/j.apenergy.2010.05.012.
37. Leung A, Chau F, Gao J. A review on applications of wavelet transform techniques in chemical analysis: 1989–1997. *Chemometrics and Intelligent Laboratory Systems* 1998; **43**(1-2): 165–184. DOI: 10.1016/S0169-7439(98)00080-X.
38. Zhu K, Wong YS, Hong GS. Wavelet analysis of sensor signals for tool condition monitoring: a review and some new results. *International Journal of Machine Tools and Manufacture* 2009; **49**(7-8): 537–553. DOI: 10.1016/j.ijmachtools.2009.02.003.
39. Domingues MO, Mendes O, da Costa AM. On wavelet techniques in atmospheric sciences. *Advances in Space Research* 2005; **35**(5): 831–842. DOI: 10.1016/j.asr.2005.02.097.
40. Szilagyi J, Parlange MB, Katul GG, Albertson JD. An objective method for determining principal time scales of coherent eddy structures using orthonormal wavelets. *Advances in Water Resources* 1999; **22**(6): 561–566. DOI: 10.1016/S0309-1708(98)00046-3.
41. Krusche N, De Oliveira A. Characterization of coherent structures in the atmospheric surface layer. *Boundary-Layer Meteorology* 2004; **110**(2): 191–211. DOI: 10.1023/A:1026096805679.
42. Watanabe T. Large-eddy simulation of coherent turbulence structures associated with scalar ramps over plant canopies. *Boundary-Layer Meteorology* 2004; **112**(2): 307–341. DOI: 10.1023/B:BOUN.0000027912.84492.54.
43. Percival DB, Walden AT. *Wavelet Methods for Time Series Analysis*. Cambridge University Press: Cambridge, 2000.
44. Mallat S. *A Wavelet Tour of Signal Processing*, (2nd edn). Academic Press: California, 1999.
45. Burrus S, Gopinath R, Guo H. *Introduction to Wavelets and Wavelet Transforms*. Prentice Hall: Englewood Cliffs, 1998.
46. Aboufadel E, Schlicker S. *Discovering Wavelets*. Wiley, John & Sons, Inc.: New York, 1999.
47. Boggess A, Narcowich FJ. *A First Course in Wavelets with Fourier Analysis*. Prentice Hall: Upper Saddle River, 2001.
48. Madsen H. *Time Series Analysis*. Chapman & Hall, CRC: Boca Raton, 2007.
49. Katzenstein W, Fertig E, Apt J. The variability of interconnected wind plants. *Energy Policy* 2010; **38**(8): 4400–4410. DOI: 10.1016/j.enpol.2010.03.069.
50. Cuerva A, Pérez C, Sanz-Andrés A, García-Barquero C, Franchini S, Zubiaur R. On the behaviour of large complex wind farms, *European Wind Energy Conference*, London, 2004.
51. Madsen H, Pinson P, Kariniotakis G, Nielsen HA, Nielsen TS. Standardizing the performance evaluation of short-term wind power prediction models. *Wind Engineering* 2005; **29**(6): 475–489. DOI: 10.1260/030952405776234599.